



Blaconá, María Teresa

Mari, Gonzalo

Méndez, Fernanda

Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística, Facultad de Ciencias Económicas y Estadística, UNR

ESTRATIFICACIÓN DE POBLACIONES ASIMÉTRICAS

1. INTRODUCCIÓN

Uno de los objetivos del muestreo en poblaciones finitas es obtener resultados que sean lo más precisos posibles. Existen técnicas que permiten lograr una reducción en el error muestral, lo cual se logra en muchos casos a partir de utilizar información auxiliar. Entre los métodos muestrales que consideran este tipo de información, uno de los más conocidos es el muestreo estratificado, que considera la partición de la población en grupos mutuamente excluyentes y homogéneos intra grupo. El muestreo estratificado considera la selección de muestras probabilísticas en cada uno de los estratos en forma independiente.

Uno de los elementos que se deben tener en cuenta a la hora de estratificar una población es la distribución de la variable bajo estudio. Existen en la literatura un conjunto de métodos de estratificación que son válidos para distribuciones simétricas, tales como Dalenius y Hodges (1959), o Ekman (1959).

El problema de éstos métodos es que no brindan soluciones precisas cuando la distribución de la variable bajo estudio es asimétrica. Existen métodos que son apropiados para este tipo de distribuciones, entre los que se pueden mencionar Hidiroglou (1986), Lavallée e Hidiroglou (1988), Kozak y Verma (2006) como métodos óptimos que trabajan además con un estrato de inclusión forzosa. Existen también métodos aproximados, como por ejemplo Gunning y Horgan (2004), que resultan ser menos precisos que los anteriores pero de una aplicación más simple.

Los trabajos existentes en la literatura evalúan métodos de estratificación aplicados a poblaciones con distintos grados de asimetría de tamaños medianos a grandes, pero en la mayoría de los casos no se presentan resultados para tamaños pequeños de población. Debido a que el objetivo es estratificar poblaciones de unidades económicas dentro de la provincia de Santa Fe, y la estratificación se realiza para dominios de estudio (agrupamientos de ramas de actividad) que contienen un número reducido de unidades económicas, se prueban las bondades de los métodos en poblaciones con dichas características y distintos nivel de asimetría.

El objetivo de este artículo es comparar la eficiencia de la estratificación geométrica propuesta por Gunning y Horgan (2004) y los dos enfoques de optimización para estratificación, Lavallée e Hidiroglou (1988), y Kozak y Verma (2006), los cuales se basan en el uso de métodos de optimización numérica. Dado que los últimos dos métodos son iterativos y pueden variar de acuerdo a la información inicial que se le suministra, se probarán también diversos



valores iniciales tal como lo propone Kozak y Verma (2006) para método descripto en el mismo trabajo y Gunning, Horgan y Keogh (2008) para el método de Lavallée e Hidiroglou.

En este artículo en la sección 2 se introduce una breve definición de los conceptos a utilizar, en la sección 3 se presentan distintos métodos de estratificación asimétrica, en la sección 4 se muestra una evaluación de los métodos realizado por simulación y en la sección 5 se desarrollan las conclusiones y futuros pasos de la investigación.

2. EL PROBLEMA DE LA ESTRATIFICACIÓN

Se quiere estratificar una población asimetría positiva U de N elementos, basados en un vector N dimensional de valores $\mathbf{x} = (x_1, \dots, x_N)^T$, conocidos previamente, de una variable de estratificación x . En el presente trabajo se considera que la variable de estratificación x es idéntica a la variable de estudio y .

El problema de la estratificación se puede ver desde dos lugares distintos. En el primero, el problema consiste en la construcción de L estratos sujeto a un tamaño de muestra dado n . En el segundo enfoque, que es el que se considera en el presente trabajo, el problema es la construcción de estratos que minimicen el tamaño de muestra de una población con respecto a un nivel de precisión en las estimaciones dado.

Sea una población U de tamaño N y se asume que la misma se encuentra estratificada de acuerdo a un método cualquiera. De esta forma, la población está compuesta por U_h estratos de tamaño N_h . En cada uno de los estratos, se selecciona una muestra simple al azar de tamaño n_h , de forma tal que $n = \sum_{h=1}^L n_h$. Se asume que se utiliza la asignación a_h para adjudicar la muestra entre los estratos, de forma tal que $n_h = na_h$. Si el tamaño de muestra es n , luego la variancia del estimador del total t_x , $\hat{t}_x = \sum_{h=1}^L \sum_{i \in S_h} (N_h/n_h) x_{hi}$, es

$$V(\hat{t}_x) = \sum_{h=1}^L N_h^2 \left(\frac{1}{na_h} - \frac{1}{N_h} \right) S_{xh}^2 \quad (2.1)$$

donde $S_{xh}^2 = \frac{1}{N_h - 1} \sum_{i \in S_h} (x_{hi} - \bar{x}_h)^2$ es la variancia poblacional de x en el estrato h ,

\bar{x}_h es la media poblacional de x en el estrato h

Se fija el coeficiente de variación (CV) igual a c . El tamaño de muestra necesario para que el estimador \hat{t}_x tenga un CV igual a c se obtiene resolviendo la ecuación $V(\hat{t}_x) = c^2 t_x^2$, donde

$t_x = \sum_{h=1}^L \sum_{i=1}^{N_h} x_{hi}$. El tamaño de muestra queda determinado por

$$n = \frac{\sum_{h=1}^L N_h^2 S_{xh}^2 / a_h}{c^2 t_x^2 + \sum_{h=1}^L N_h S_{xh}^2} \quad (2.2)$$

Cabe mencionar que en el presente trabajo se considera la adjudicación de Neyman sobre



la variable de estratificación x , siendo $a_h = \frac{N_h S_{xh}}{\sum_{h=1}^L N_h S_{xh}}$.

3. MÉTODOS DE ESTRATIFICACIÓN

3.1. Método Geométrico (GEO)

Gunning y Horgan (2004) desarrollaron un método aproximado para estratificar poblaciones asimétricas. Esta estrategia recurre a una progresión geométrica para determinar los límites de los estratos, bajo el supuesto de obtener coeficientes de variación (CV) aproximadamente iguales en cada uno de los estratos. La misma se basa en una observación de Cochran (1961), que con límites de estratos cercanos a los óptimos, los coeficientes de variación resultan aproximadamente iguales en todos los estratos.

El objetivo al estratificar una población es subdividirla en intervalos, siendo los límites de éstos $k_0 < k_1 < \dots < k_L$. Dado que no se cuenta con la variable bajo estudio, y , se utiliza una variable auxiliar conocida x que esté correlacionada con la variable de interés.

Estos límites se obtienen teniendo como objetivo que los coeficientes de variación $CV_h = S_{xh}^2 / \bar{x}_h$ sean iguales para todos los estratos $h=1, \dots, L$, o sea,

$$\frac{S_{x1}}{\bar{x}_1} = \frac{S_{x2}}{\bar{x}_2} = \dots = \frac{S_{xL}}{\bar{x}_L}$$

Bajo el supuesto que la distribución dentro de cada estrato es aproximadamente uniforme, tenemos que

$$\bar{x}_h \approx \frac{k_h + k_{h-1}}{2}$$

$$S_{xh} \approx \frac{1}{\sqrt{12}}(k_h - k_{h-1})$$

Luego, una aproximación a los coeficientes de variación viene dada por

$$CV_h = \frac{(k_h - k_{h-1}) / \sqrt{12}}{(k_h + k_{h-1}) / 2}$$

con lo cual para obtener coeficientes de variación iguales se debe cumplir

$$\frac{k_{h+1} - k_h}{k_{h+1} + k_h} = \frac{k_h - k_{h-1}}{k_h + k_{h-1}}$$

Esta relación recurrente se reduce a

$$k_h^2 = k_{h+1} k_{h-1}$$

siendo los límites de los estratos los términos de una progresión geométrica



$$k_h = ar^h \quad (h = 0, 1, \dots, L)$$

Luego $a=k_0$, el valor mínimo de la variable, y $ar^L=k_L$, el máximo de la variable, con lo cual la razón constante puede calcularse como $r=(k_L/k_0)^{1/L}$.

3.2. Método de Lavallée-Hidiroglou (LH)

Lavallée e Hidiroglou (1988) presentaron un procedimiento iterativo para estratificar poblaciones asimétricas en un estrato de inclusión forzosa, o sea donde todas las unidades del mismo son examinadas, y un número de estratos de los cuales se seleccionan muestras. Los límites de los estratos se obtienen utilizando una variable auxiliar conocida correlacionada con la variable bajo estudio. El algoritmo comienza con un conjunto de límites iniciales y reemplaza los mismos iterativamente, utilizando el procedimiento de Sethi (1963), hasta obtener límites que minimicen el tamaño de muestra sujeto a un nivel de precisión fijo.

El algoritmo ordena los datos en orden ascendente y obtiene L estratos definidos por los límites $k_0 < k_1 < \dots < k_{L-1} < k_L$, donde $k_0 = \min(x)$ y $k_L = \max(x)$. El límite k_{L-1} crea un estrato donde todas las N_L unidades del mismo son consideradas, o sea, un estrato de inclusión forzosa. El resto de las $n - N_L$ unidades son adjudicadas en los $L-1$ estratos. Resulta obvio resaltar que $n > N_L$.

El objetivo es elegir límites que minimicen el tamaño de muestra n para un nivel de precisión dado, generalmente el coeficiente de variación del estimador de la media,

$$cv(\hat{\bar{x}}_{st}) = \sqrt{V(\hat{\bar{x}}_{st})} / \bar{x} = c.$$

Dado el estrato de inclusión forzosa, la variancia resulta

$$V(\hat{\bar{x}}_{st}) = \sum_{h=1}^{L-1} W_h^2 \left(1 - \frac{n_h}{N_h} \right) \frac{S_{xh}^2}{n_h} \quad (3.2.1)$$

y considerando $a_h = n_h / (n - N_L)$, la proporción de las $n - N_L$ unidades adjudicadas al estrato h ($h=1, \dots, L-1$), (3.2.1) puede ser escrito como

$$V(\hat{\bar{x}}_{st}) = \sum_{h=1}^{L-1} W_h^2 \left(1 - \frac{(n - N_L)a_h}{N_h} \right) \frac{S_{xh}^2}{(n - N_L)a_h} \quad (3.2.2)$$

Resolviendo (3.2.2) para el tamaño de muestra n se obtiene

$$n = N_L + \frac{\sum_{h=1}^{L-1} W_h^2 S_{xh}^2 / a_h}{V(\hat{\bar{x}}_{st}) + \sum_{h=1}^{L-1} W_h^2 S_{xh}^2 / N} \quad (3.2.3)$$

y escribiendo $V(\hat{\bar{x}}_{st}) = \bar{x}^2 c^2$ se obtiene

$$n = N_L + \frac{\sum_{h=1}^{L-1} W_h^2 S_{xh}^2 / a_h}{\bar{x}^2 c^2 + \sum_{h=1}^{L-1} W_h^2 S_{xh}^2 / N} \quad (3.2.4)$$

Lavallée e Hidiroglou (1988) sugirieron que en (3.2.4), n sea tratado como una función de los límites de los estratos k_1, k_2, \dots, k_{L-1} . Luego, los límites óptimos, o sea aquellos que minimicen n sujeto a un CV fijo c , se obtienen igualando a 0 las derivadas parciales de (3.2.4) con res-



pecto a k_h , o sea,

$$\frac{\partial n}{\partial k_1} = \frac{\partial n}{\partial k_2} = \dots = \frac{\partial n}{\partial k_{L-1}} = 0 \quad (3.2.5)$$

De (3.2.5) se obtiene una serie que se asemeja a una ecuación cuadrática para los límites k_h

$$\alpha_h k_h^2 + \beta_h k_h + \gamma_h = 0, \quad 1 \leq h \leq L-1 \quad (3.2.6)$$

Debido a que los coeficientes $\alpha_h, \beta_h, \gamma_h$ son funciones de W_h, S_{xh}, \bar{x}_h y ah , que a su vez depende de k_h, k_{h-1} y k_{h+1} , la solución a las ecuaciones (3.2.6) no es trivial. Lavallée e Hidiroglou (1988) proveen una solución a partir del siguiente procedimiento iterativo:

- i) Ordenar la población en orden ascendente
- ii) Comenzar con límites iniciales arbitrarios $k'_1 < k'_2 < \dots < k'_{L-1}$. En el presente trabajo se consideran para este método límites que forman grupos iguales en la población.
- iii) Basado en esos límites, calcular W'_h, S'_{xh}, x'_h para cada estrato $h=1, \dots, L-1$
- iv) Calcular n utilizando la fórmula (3.2.4)
- v) Las N_L unidades del estrato de inclusión forzosa son seleccionadas en la muestra, y el resto de las $n-N_L$ unidades muestrales son seleccionadas de los restantes $L-1$ estratos utilizando un método de adjudicación apropiado. En el presente trabajo se considera la adjudicación de Neyman.
- vi) Reemplazar el conjunto inicial de límites tomando la raíz mayor de (3.2.6), o sea,

$$k''_h = \frac{-\beta'_h \pm \sqrt{(\beta'_h)^2 - 4\alpha'_h \gamma'_h}}{2\alpha'_h}$$

- vii) Repetir pasos iii) a vi) con el nuevo conjunto de límites, continuando hasta que dos conjuntos sucesivos sean idénticos o con diferencias despreciables.

3.3. Método de Lavallée-Hidiroglou Geométrico (LHG)

Gunning, Horgan y Keogh (2008) presentaron una modificación para el método de Lavallée-Hidiroglou descrito en el punto 3.2. Entre los múltiples inconvenientes que pueden presentarse a la hora de aplicar el método de LH, se pueden mencionar: los límites finales dependen del conjunto de límites iniciales ocurriendo que el tamaño de muestra mínimo obtenido sea local y no un mínimo global, convergencia lenta, no convergencia al tamaño de muestra mínimo verdadero. La propuesta de estos autores es una estrategia para improvisar la convergencia del algoritmo de LH.

El algoritmo es similar al presentado en el punto 3.2, con la diferencia que en el punto ii) se escogen los límites de estratos obtenidos a partir de la estratificación geométrica descrita en el punto 3.1.

Los autores demuestran empíricamente que el procedimiento LHG converge más rápidamente a los límites óptimos, y que en muchos casos resulta en un tamaño de muestra menor que los obtenidos con el algoritmo original.



3.4. Método de Kozak (K)

Kozak (2004) desarrolla un método de estratificación que brinda límites de estratificación cercanos al óptimo cuando se cuenta con variables de estratificación con distribución asimétrica. Está basado en un método de optimización numérica que brinda límites de estratos. Al igual que el método LH, posee el beneficio de brindar un estrato de inclusión forzosa y adjudicar el tamaño de muestra restante entre los $L-1$ estratos donde se seleccionará muestra.

El método consiste en encontrar límites de estratos $k_1 < k_2 < \dots < k_L$, que minimicen la función objetivo dada por

$$n = N_L + \frac{\sum_{h=1}^{L-1} W_h^2 S_{xh}^2 / a_h}{\bar{x}^2 c^2 + \sum_{h=1}^{L-1} W_h S_{xh}^2 / N} \quad (3.4.1)$$

donde n es el tamaño de muestra mínimo requerido para obtener el nivel de precisión c , bajo las restricciones

$$\begin{aligned} N_h &\geq 2, \quad h = 1, \dots, L \\ 2 &\leq n_h \leq N_h, \quad h = 1, \dots, L \end{aligned} \quad (3.4.2)$$

En el presente trabajo, en la función objetivo (3.4.1) se considera una asignación de Neyman en los $L-1$ estratos donde se selecciona muestra. Por otra parte, y tal como pasa con los métodos anteriores, se considera una variable de estratificación x , conocida, y correlacionada con la variable bajo estudio y .

El algoritmo propuesto por Kozak (2004), es el método de búsqueda aleatoria y resuelve el problema de estratificación óptima minimizando la función (3.4.1) sujeto a las restricciones (3.4.2):

- i) Ordenar la población por los valores de la variable de estratificación
- ii) Elegir un conjunto de límites de estratos iniciales. Pueden ser enteros aleatorios o límites obtenidos a partir de otros métodos clásicos. Obtener las ubicaciones de los límites en la muestra ordenada. Sea el vector de posiciones $\mathbf{j} = (j_1, j_2, \dots, j_{L-1})$. Calcular la función (3.4.1) para el vector \mathbf{j} , $n = n(\mathbf{j})$
- iii) Para $r=0, 1, \dots, R$ repetir los siguientes pasos:

- a) Generar un vector \mathbf{j}' seleccionando al azar una posición j_i y cambiándolo por

$$\begin{aligned} j'_i &= j_i + k \\ j'_l &= j_l \quad \text{para } l = 1, \dots, L-1, \quad l \neq i \end{aligned} \quad (3.4.3)$$

donde k es un entero aleatorio, $k \in \langle -p; 1 \rangle \cup \langle 1; p \rangle$; siendo p un entero fijo de acuerdo al tamaño de la población.

- b) Calcular el valor de la función $n' = n(\mathbf{j}')$
- c) Si las restricciones (3.4.2) se cumplen y $n(\mathbf{j}') \leq n(\mathbf{j})$, aceptar $\mathbf{j}_{r+1} = \mathbf{j}'$ (donde \mathbf{j}_{r+1} es el vector de posiciones en la próxima iteración), sino $\mathbf{j}_{r+1} = \mathbf{j}$

iv) Finalizar el algoritmo si $r=R$ o si en los últimos m pasos el tamaño de muestra no disminuyó. Obtener los límites de los estratos en base a las posiciones finales \mathbf{j} .



Una de las cuestiones a tener en cuenta en el presente método es la elección de los límites iniciales en el algoritmo de optimización. Una de las opciones es considerar límites arbitrarios, en este caso aquellos que formen grupos iguales en la población (K). Una segunda opción es la de considerar límites que provengan de métodos clásicos de estratificación. Para este trabajo se evalúan los límites iniciales que provienen de la estratificación geométrica (KG).

4. EVALUACIÓN DE LOS MÉTODOS

Para evaluar el comportamiento de los métodos de estratificación se utilizan poblaciones simuladas de distintos tamaños y niveles de asimetría. Como ya se mencionó, los tamaños de las poblaciones serán pequeños y medianos (200, 250, 333, 500 y 1000) para representar las poblaciones económicas que se presentan en distintas ramas o grupos de ramas de actividad en la provincia de Santa Fe. Cabe destacar que en la bibliografía existente no es usual estudiar estos procedimientos para tamaños de población tan reducidos.

Para cada uno de los tamaños de población se simularon poblaciones con cuatro grados distintos de asimetría. Las características de las poblaciones se presentan en el Cuadro 4.1, como así también los histogramas de las mismas en los Gráfico 4.1. a 4.5.

Cuadro 4.1. Características de las poblaciones simuladas

Nivel Asimetría	Tamaño	Mínimo	Máximo	Media	Variación	Asimetría
1	1,000	104	904	297.0	1.7E+04	1.0
1	500	107	904	297.0	1.7E+04	1.0
1	333	110	794	297.0	1.7E+04	0.9
1	250	110	794	297.0	1.7E+04	1.0
1	200	111	794	297.0	1.7E+04	1.0
2	1,000	101	7,310	1,097	9.8E+05	1.9
2	500	101	6,167	1,094	9.6E+05	1.8
2	333	102	6,167	1,097	9.7E+05	1.8
2	250	102	6,167	1,100	9.8E+05	1.8
2	200	102	6,167	1,103	9.9E+05	1.8
3	1,000	100	51,571	2,641	3.0E+07	4.5
3	500	100	51,571	2,674	3.2E+07	4.6
3	333	100	51,571	2,697	3.3E+07	4.7
3	250	100	51,571	2,733	3.5E+07	4.8
3	200	100	51,571	2,754	3.6E+07	4.9
4	1,000	100	81,445	1,891	2.6E+07	7.9
4	500	100	75,680	1,875	2.5E+07	7.9
4	333	100	69,914	1,917	2.8E+07	7.7
4	250	100	69,914	1,972	3.2E+07	7.9
4	200	100	69,914	2,033	3.7E+07	7.8



Gráfico 4.1 Histograma de Poblaciones de tamaño 1000 y 4 niveles de asimetría

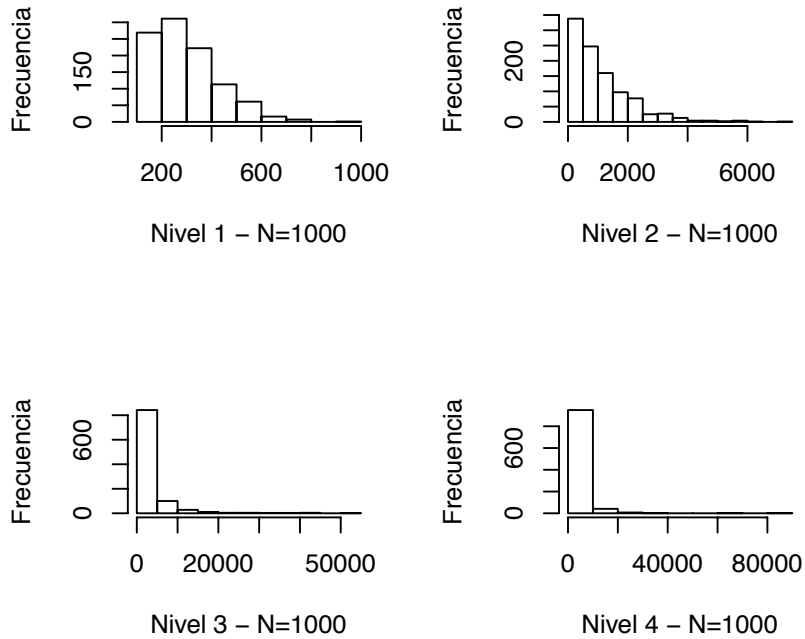


Gráfico 4.2 Histograma de Poblaciones de tamaño 500 y 4 niveles de asimetría

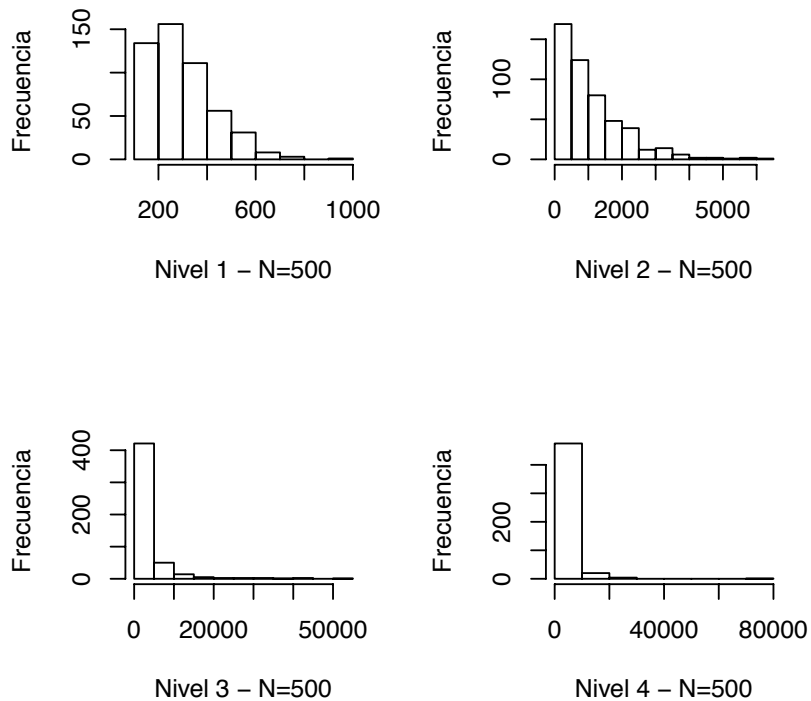




Gráfico 4.3 Histograma de Poblaciones de tamaño 333 y 4 niveles de asimetría

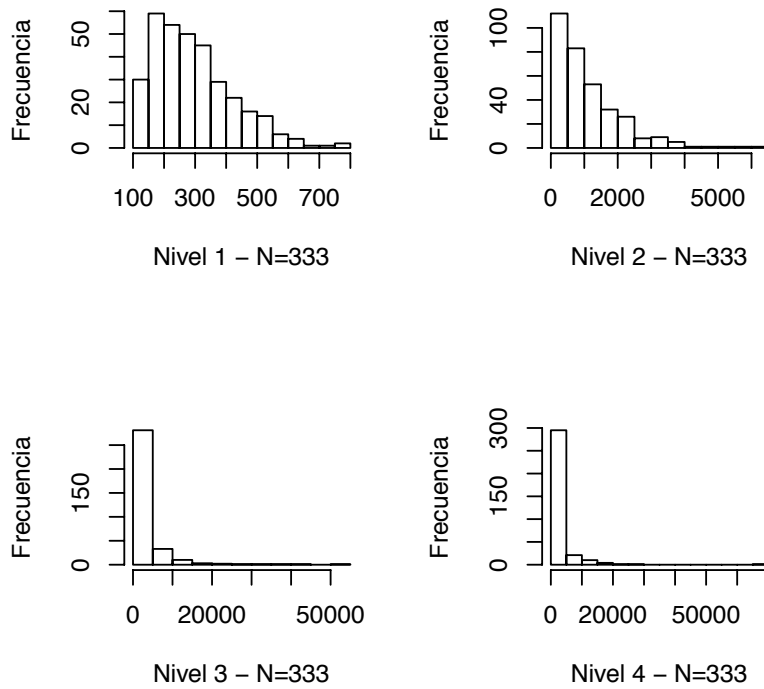


Gráfico 4.4 Histograma de Poblaciones de tamaño 250 y 4 niveles de asimetría

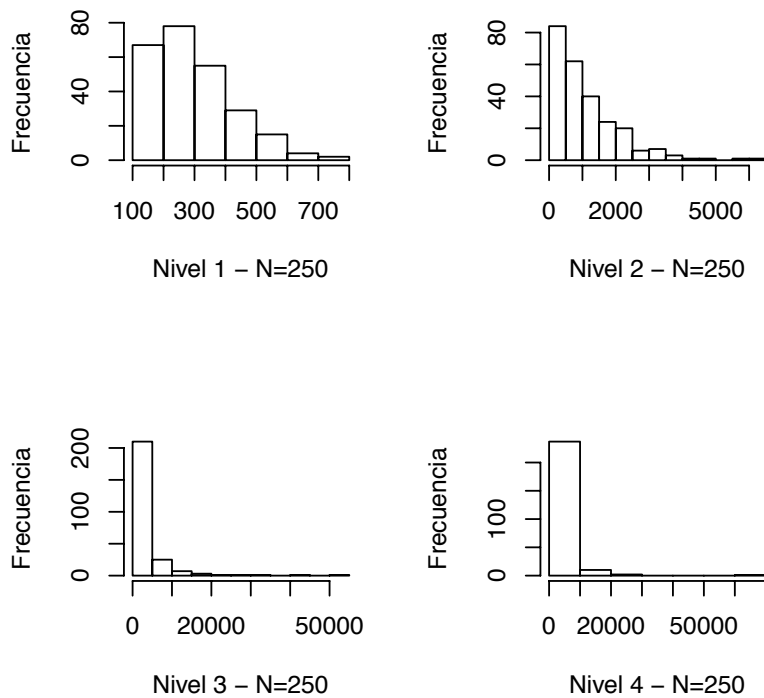
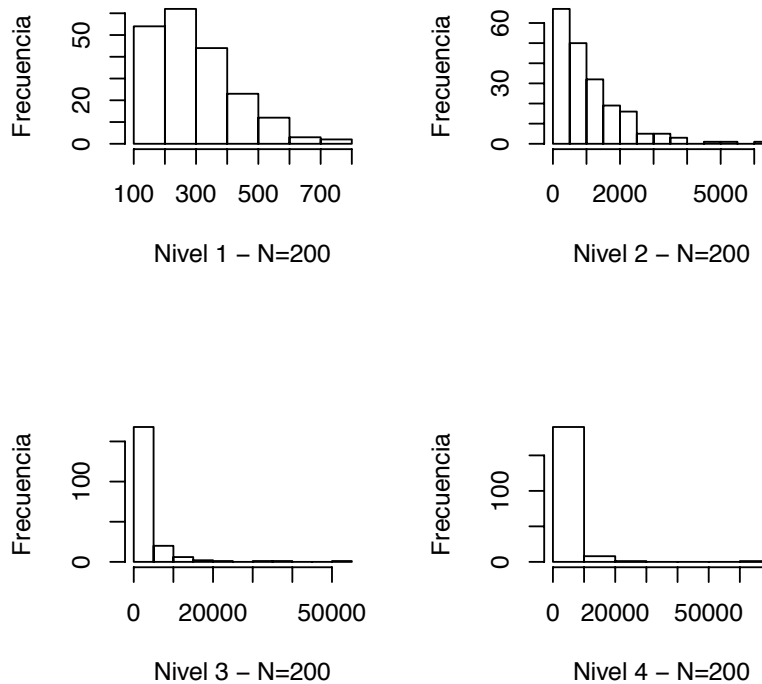




Gráfico 4.5 Histograma de Poblaciones de tamaño 200 y 4 niveles de asimetría



Las poblaciones se estratifican de acuerdo a las metodologías descritas en el punto 3, o sea, método geométrico (GEO), método de Lavallée-Hidiroglou (LH), método de Lavallée-Hidiroglou Geométrico (LHG), y el método de Kozak con sus dos variantes: con límites iniciales arbitrarios (K) y con límites iniciales geométricos (KG). Con cada uno de los métodos se obtuvo el tamaño de muestra mínimo para obtener 4 niveles distintos de precisión ($CV=0.10, 0.05, 0.03, \text{ y } 0.01$) para la estimación de la media considerando 3, 4 y 5 estratos.

Para evaluar la eficiencia de dos métodos se considera la siguiente medida relativa

$$efrel_{i,j} = \frac{n_i(cv)}{n_j(cv)} \tag{4.1}$$

donde i y j son los métodos de estratificación a utilizar ($i, j= \text{GEO, LH, LHG, K, KG}$), y $n_i(cv)$ y $n_j(cv)$ son los tamaños de muestra mínimos requeridos para obtener el nivel de precisión deseado cv bajo los métodos i y j respectivamente. Debido a que los tamaños de muestras de los métodos LH y K son similares si se trabaja con límites iniciales arbitrarios o geométricos, se optó por la última opción a la hora de calcular las eficiencias relativas. Esta decisión está sustentada en el hecho que se observó en algunos casos que el método LH convergía con límites iniciales geométricos cuando no lo hacía con los arbitrarios.

Se presenta a modo de ejemplo el cuadro 4.2 donde se evalúan los métodos para poblaciones de tamaño 1000 y cuando se desea un CV igual a 0.01. El resto de los cuadros se presentan en el Anexo.



Cuadro 4.2. Tamaños de muestras mínimos para obtener un $CV=0.01$ en la estimación de la media en poblaciones de tamaño 1000 de distintivos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	243	278	278	277	277	0.87	0.88	1.00
4	179	181	181	180	180	0.99	0.99	1.01
5	126	121	124	.	121	1.02	1.04	1.02
Nivel de Asimetría 2								
3	452	396	396	395	395	1.14	1.14	1.00
4	360	300	300	299	299	1.20	1.20	1.00
5	291	224	224	222	222	1.30	1.31	1.01
Nivel de Asimetría 3								
3	365	298	298	298	297	1.22	1.23	1.00
4	298	242	242	239	239	1.23	1.25	1.01
5	244	194	195	.	193	1.25	1.26	1.01
Nivel de Asimetría 4								
3	327	230	230	230	230	1.42	1.42	1.00
4	233	190	230	188	189	1.01	1.23	1.22
5	210	168	190	166	158	1.11	1.33	1.20

Conclusiones

- La utilización de valores iniciales de límites de estratos geométricos en K no provocó una disminución en el tamaño de muestra obtenido
- Utilizar límites iniciales geométricos en LH solucionó problemas de convergencia, en particular para: Nivel 1, $CV=0.01$ y $N=1000$, 500 y 333
- El método LH proporciona iguales tamaños de muestra que el método K, resultando bajo algunos escenarios mejor en alrededor de un 5%
- Los métodos óptimos logran un tamaño de muestra menor que el geométrico comenzando con ganancias en promedio del 8% para el Nivel 1, llegando a una disminución del 30% en el tamaño de muestra para el Nivel 4 de asimetría. Esta mejora es más marcada en poblaciones más asimétricas y una precisión alta ($cv=0.01$)
- El método geométrico muestra mejores resultados en poblaciones de asimetrías leves o moderada y esta mejora se incrementa cuando se requiere una precisión inferior.



- El método KG presenta resultados incoherentes respecto al método K en población con nivel 4 de asimetría y alta precisión, resultando ser mayor el tamaño de muestra. Esto se debe a que el método para esa población considera los 2 estratos más altos como inclusión forzosa para poder alcanzar el nivel de precisión de $cv=0.01$. Esta situación no se presenta con límites iniciales arbitrarios.
- Cuando se fija el cv en 0.10, las restricciones respecto al tamaño de muestra en cada estrato para el método K y KG no se cumplen. Se observa en las tablas correspondientes a ese cv que en algunos estratos el tamaño de muestra es inferior a 2 (tamaño de muestra total inferior a 2L). Esto se debe a que en todos los casos con esos tamaños de muestra la precisión resulta superior a la requerida, por lo que no se necesita que la restricción sea satisfecha.

5. CONCLUSIONES Y ESTUDIOS FUTUROS

El problema a la hora de realizar encuestas por muestreo donde se requiera el uso de muestreo estratificado en poblaciones asimétricas recibe un trato diferencial en la literatura debido a que no pueden aplicarse los métodos clásicos de estratificación. Por otra parte, es muy difícil encontrar compensaciones de los métodos apropiados para poblaciones de tamaños medianas a pequeñas.

En el presente estudio se trabajó con una serie de métodos que son apropiados para estratificar poblaciones asimétricas, y se evaluaron las propiedades de los mismos bajo diversos escenarios de nivel de asimetría, tamaño de población, número de estratos. En particular se consideró el tamaño de muestra mínimo necesario para lograr un nivel de precisión fijado de antemano.

Entre los métodos evaluados, el geométrico resulta ser el de mayor facilidad de aplicación, debido a que no usa ningún proceso iterativo, como si lo hacen los métodos de Kozak y Lavallée-Hidiroglou, y se basa en una proyección geométrica que resulta de fácil aplicación. El problema que tiene este método es que es aproximado, con lo cual los resultados no siempre serán cercanos al óptimo. De todas formas, se observó que para poblaciones con una asimetría leve a moderada y con pocos requerimientos de precisión obtuvo mejores resultados que sus competidores.

El método de Kozak es un método de los llamados óptimos, dado que en la mayoría de los casos el se obtiene el mínimo tamaño de muestra. Utilizar límites iniciales geométricos o arbitrarios no mostró para las poblaciones y escenarios propuestos ninguna diferencia, por lo tanto se podrá considerar el uso de iniciales arbitrarios a la hora de aplicarlo para que sea más simple. En poblaciones con un mayor nivel de asimetría y alta precisión pretendida se logran resultados visiblemente mejores a los que se obtienen con la estratificación geométrica.

El método de Lavallée-Hidiroglou también es uno de los llamados óptimos. En este caso, utilizar los límites geométricos solucionó en algunos escenarios el problema de convergencia que se observaba cuando se usaban los límites arbitrarios, pero no produjo disminuciones importantes en los tamaños de muestra requeridos respecto al método con límites arbitrarios. Respecto al método de Kozak, no se observan diferencias importantes, si bien el método de Lavallée-Hidiroglou resultó igual o mejor.

Una de las cuestiones a considerar en estudios futuros es la discrepancia que existe entre la variable de estratificación y la variable bajo estudio. En el presente trabajo se las consideró iguales, pero en la práctica no es lo que sucede. Los métodos analizados en este trabajo permiten considerar estas diferencias a la hora de realizar el proceso de estratificación



6. BIBLIOGRAFÍA

- Baillargeon, S., Rivest, L.P. (en Prensa). A General Algorithm for Univariate Stratification. *International Statistical Review*.
http://www.mat.ulaval.ca/fileadmin/Pages_personnelles_des_profs/lpr/Publications/95-Baillargeon_Rivest.pdf
- Cochan, W.G. (1961). Comparison of methods for determining stratum boundaries. *Bulletin of the International Statistical Institute*, 32, 345-358.
- Dalenius, T., Hodges, J.L. (1959). Minimum variance stratification. *Journal of the American Statistical Association*, 54, 88-101.
- Gunning, P. and Horgan, J. M. (2004). A simple algorithm for stratifying skewed populations. *Survey Methodology*, 30, 159-166
- Gunning, P., Horgan, J.M., Keogh, G. (2008). An Implementation Strategy for Efficient Convergence of the Lavallée and Hidiroglou Stratification Algorithm. *Journal of Official Statistics*, 24, 213-228.
- Hidiroglou, M. A. (1986). The construction of a self-representing stratum of large units in survey design. *The American Statistician*, 40, 27-31
- Kozak, M. (2004). Optimal Stratification using Random Search Method in Agricultural Surveys. *Statistics in Transition*, 6, 797-806.
- Kozak, M., Verma, M.R. (2006). Geometric Versus Optimization Approach to Stratification: A Comparison of Efficiency. *Survey Methodology*, 32, 157-163.
- Lavallée, P, Hidiroglou, M. (1988). On the stratification of skewed populations. *Survey Methodology*, 14, 33-43.
- Särndal C. E., Swensson B. and Wretman J. (1992). *Model Assisted Survey Sampling*. Springer Verlag: New York.
- Sethi, V. K. (1963). A note on the optimum stratification of populations for estimating the population means. *Australian Journal of Statistics*, 5, 20-33



7. ANEXO

Cuadro A.3. Tamaños de muestras mínimos para obtener un $CV=0.01$ en la estimación de la media en poblaciones de tamaño 200 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	113	103	103	103	102	1.10	1.11	1.01
4	93	81	81	79	79	1.15	1.18	1.03
5	76	67	68	65	64	1.12	1.19	1.06
Nivel de Asimetría 2								
3	136	118	118	116	116	1.15	1.17	1.02
4	118	98	98	97	97	1.20	1.22	1.01
5	107	86	84	83	83	1.27	1.29	1.01
Nivel de Asimetría 3								
3	102	82	82	81	81	1.24	1.26	1.01
4	84	69	69	69	68	1.22	1.24	1.01
5	75	61	61	58	60	1.23	1.25	1.02
Nivel de Asimetría 4								
3	87	62	78	62	61	1.12	1.43	1.28
4	71	53	62	51	53	1.15	1.34	1.17
5	62	46	62	44	49	1.00	1.27	1.27



Cuadro A.4. Tamaños de muestras mínimos para obtener un $CV=0.01$ en la estimación de la media en poblaciones de tamaño 250 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	130	120	120	119	119	1.08	1.09	1.01
4	103	92	92	91	90	1.12	1.14	1.02
5	82	75	75	74	72	1.09	1.14	1.04
Nivel de Asimetría 2								
3	163	141	141	140	140	1.16	1.16	1.01
4	140	116	116	115	114	1.21	1.23	1.02
5	126	100	100	97	97	1.26	1.30	1.03
Nivel de Asimetría 3								
3	125	99	99	97	98	1.26	1.28	1.01
4	101	82	83	83	82	1.22	1.23	1.01
5	90	72	72	70	71	1.25	1.27	1.01
Nivel de Asimetría 4								
3	106	75	95	75	74	1.12	1.43	1.28
4	84	64	75	63	64	1.12	1.31	1.17
5	76	54	64	54	57	1.19	1.33	1.12



Cuadro A.5. Tamaños de muestras mínimos para obtener un $CV=0.01$ en la estimación de la media en poblaciones de tamaño 333 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	151	146	146	144	144	1.03	1.05	1.01
4	118	109	109	107	107	1.08	1.10	1.02
5	91	84	84	.	83	1.08	1.10	1.01
Nivel de Asimetría 2								
3	202	177	177	176	176	1.14	1.15	1.01
4	172	144	144	141	142	1.19	1.21	1.01
5	150	120	121	120	119	1.24	1.26	1.02
Nivel de Asimetría 3								
3	159	125	125	124	124	1.27	1.28	1.01
4	126	103	103	102	101	1.22	1.25	1.02
5	109	90	90	89	92	1.21	1.18	0.98
Nivel de Asimetría 4								
3	132	96	124	95	95	1.06	1.39	1.31
4	105	81	96	79	80	1.09	1.31	1.20
5	95	70	81	68	68	1.17	1.40	1.19



Cuadro A.5. Tamaños de muestras mínimos para obtener un $CV=0.01$ en la estimación de la media en poblaciones de tamaño 500 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	193	188	188	187	187	1.03	1.03	1.01
4	147	136	136	135	135	1.08	1.09	1.01
5	112	100	100	.	98	1.12	1.14	1.02
Nivel de Asimetría 2								
3	272	243	243	242	242	1.12	1.12	1.00
4	227	192	192	191	191	1.18	1.19	1.01
5	193	155	155	154	153	1.25	1.26	1.01
Nivel de Asimetría 3								
3	219	173	173	172	172	1.27	1.27	1.01
4	176	139	139	139	139	1.27	1.27	1.00
5	148	126	126	123	121	1.17	1.22	1.04
Nivel de Asimetría 4								
3	186	135	179	133	133	1.04	1.40	1.35
4	141	113	135	111	111	1.04	1.27	1.22
5	130	96	112	95	100	1.16	1.30	1.12



Cuadro A.6. Tamaños de muestras mínimos para obtener un $CV=0.03$ en la estimación de la media en poblaciones de tamaño 200 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	30	41	41	40	40	0.73	0.75	1.03
4	21	25	25	24	24	0.84	0.88	1.04
5	15	17	17	15	15	0.88	1.00	1.13
Nivel de Asimetría 2								
3	72	65	65	64	64	1.11	1.13	1.02
4	53	46	46	46	46	1.15	1.15	1.00
5	41	35	35	33	33	1.17	1.24	1.06
Nivel de Asimetría 3								
3	59	52	52	51	52	1.13	1.13	1.00
4	49	41	41	40	40	1.20	1.23	1.03
5	39	31	31	32	30	1.26	1.30	1.03
Nivel de Asimetría 4								
3	52	40	40	39	39	1.30	1.33	1.03
4	40	34	33	32	31	1.21	1.29	1.06
5	35	27	27	26	25	1.30	1.40	1.08



Cuadro A.7. Tamaños de muestras mínimos para obtener un $CV=0.03$ en la estimación de la media en poblaciones de tamaño 250 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	31	45	45	44	44	0.69	0.70	1.02
4	21	26	26	24	24	0.81	0.88	1.08
5	15	17	17	16	16	0.88	0.94	1.06
Nivel de Asimetría 2								
3	82	74	76	74	74	1.08	1.11	1.03
4	59	51	51	50	50	1.16	1.18	1.02
5	45	39	39	37	37	1.15	1.22	1.05
Nivel de Asimetría 3								
3	68	62	62	61	61	1.10	1.11	1.02
4	56	47	47	46	46	1.19	1.22	1.02
5	44	36	36	35	35	1.22	1.26	1.03
Nivel de Asimetría 4								
3	62	48	48	48	47	1.29	1.32	1.02
4	47	38	39	37	36	1.21	1.31	1.08
5	40	31	31	29	30	1.29	1.33	1.03



Cuadro A.8. Tamaños de muestras mínimos para obtener un $CV=0.03$ en la estimación de la media en poblaciones de tamaño 333 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	32	49	49	48	48	0.65	0.67	1.02
4	21	28	28	26	26	0.75	0.81	1.08
5	16	18	18	16	16	0.89	1.00	1.13
Nivel de Asimetría 2								
3	92	87	88	87	87	1.05	1.06	1.01
4	64	58	58	56	56	1.10	1.14	1.04
5	49	44	44	42	42	1.11	1.17	1.05
Nivel de Asimetría 3								
3	82	77	77	76	76	1.06	1.08	1.01
4	66	56	56	53	53	1.18	1.25	1.06
5	51	44	44	42	41	1.16	1.24	1.07
Nivel de Asimetría 4								
3	76	61	61	60	60	1.25	1.27	1.02
4	57	47	47	45	45	1.21	1.27	1.04
5	49	39	40	37	37	1.23	1.32	1.08



Cuadro A.9. Tamaños de muestras mínimos para obtener un $CV=0.03$ en la estimación de la media en poblaciones de tamaño 500 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	36	54	54	54	54	0.67	0.67	1.00
4	25	30	30	28	28	0.83	0.89	1.07
5	18	19	19	18	18	0.95	1.00	1.06
Nivel de Asimetría 2								
3	107	107	107	106	106	1.00	1.01	1.01
4	73	69	69	68	68	1.06	1.07	1.01
5	53	50	50	51	51	1.06	1.04	0.98
Nivel de Asimetría 3								
3	107	102	102	102	102	1.05	1.05	1.00
4	81	69	69	68	68	1.17	1.19	1.01
5	62	53	53	52	52	1.17	1.19	1.02
Nivel de Asimetría 4								
3	99	84	84	83	83	1.18	1.19	1.01
4	75	61	61	59	59	1.23	1.27	1.03
5	61	51	49	49	47	1.24	1.30	1.04



Cuadro A.10. Tamaños de muestras mínimos para obtener un $CV=0.03$ en la estimación de la media en poblaciones de tamaño 1000 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	36	61	61	60	60	0.59	0.60	1.02
4	27	31	31	30	30	0.87	0.90	1.03
5	18	21	21	18	18	0.86	1.00	1.17
Nivel de Asimetría 2								
3	130	146	146	145	145	0.89	0.90	1.01
4	84	87	87	85	85	0.97	0.99	1.02
5	59	59	57	59	59	1.04	1.00	0.97
Nivel de Asimetría 3								
3	172	159	159	158	158	1.08	1.09	1.01
4	118	99	99	98	98	1.19	1.20	1.01
5	85	75	75	74	74	1.13	1.15	1.01
Nivel de Asimetría 4								
3	149	141	141	140	139	1.06	1.07	1.01
4	106	92	92	90	90	1.15	1.18	1.02
5	85	76	65	74	63	1.31	1.35	1.03



Cuadro A.11. Tamaños de muestras mínimos para obtener un $CV=0.05$ en la estimación de la media en poblaciones de tamaño 200 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	13	21	21	20	20	0.62	0.65	1.05
4	10	12	12	11	11	0.83	0.91	1.09
5	7	10	10	6	6	0.70	1.17	1.67
Nivel de Asimetría 2								
3	42	41	41	40	40	1.02	1.05	1.03
4	28	27	27	25	25	1.04	1.12	1.08
5	22	20	20	19	19	1.10	1.16	1.05
Nivel de Asimetría 3								
3	41	40	40	38	38	1.03	1.08	1.05
4	32	26	26	25	25	1.23	1.28	1.04
5	24	21	21	21	20	1.14	1.20	1.05
Nivel de Asimetría 4								
3	35	33	33	32	32	1.06	1.09	1.03
4	28	23	23	22	22	1.22	1.27	1.05
5	24	20	18	18	18	1.33	1.33	1.00



Cuadro A.12. Tamaños de muestras mínimos para obtener un $CV=0.05$ en la estimación de la media en poblaciones de tamaño 250 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	13	21	21	21	21	0.62	0.62	1.00
4	10	12	12	11	11	0.83	0.91	1.09
5	7	10	10	6	6	0.70	1.17	1.67
Nivel de Asimetría 2								
3	44	46	46	45	45	0.96	0.98	1.02
4	29	30	30	27	27	0.97	1.07	1.11
5	22	21	21	18	19	1.05	1.16	1.11
Nivel de Asimetría 3								
3	48	46	46	45	45	1.04	1.07	1.02
4	35	30	30	29	29	1.17	1.21	1.03
5	27	23	23	22	22	1.17	1.23	1.05
Nivel de Asimetría 4								
3	42	39	39	38	38	1.08	1.11	1.03
4	31	28	28	26	25	1.11	1.24	1.12
5	26	23	23	20	20	1.13	1.30	1.15



Cuadro A.13. Tamaños de muestras mínimos para obtener un $CV=0.05$ en la estimación de la media en poblaciones de tamaño 333 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	13	22	22	22	22	0.59	0.59	1.00
4	10	12	12	12	12	0.83	0.83	1.00
5	7	10	10	7	7	0.70	1.00	1.43
Nivel de Asimetría 2								
3	47	51	51	50	50	0.92	0.94	1.02
4	29	32	32	30	30	0.91	0.97	1.07
5	22	22	22	20	20	1.00	1.10	1.10
Nivel de Asimetría 3								
3	60	55	55	54	54	1.09	1.11	1.02
4	41	35	35	33	33	1.17	1.24	1.06
5	30	27	27	26	26	1.11	1.15	1.04
Nivel de Asimetría 4								
3	50	48	48	47	48	1.04	1.04	1.00
4	36	32	32	31	31	1.13	1.16	1.03
5	30	27	27	25	22	1.11	1.36	1.23



Cuadro A.14. Tamaños de muestras mínimos para obtener un $CV=0.05$ en la estimación de la media en poblaciones de tamaño 500 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	14	24	24	22	22	0.58	0.64	1.09
4	10	14	13	12	12	0.77	0.83	1.08
5	9	10	10	7	7	0.90	1.29	1.43
Nivel de Asimetría 2								
3	50	60	60	60	60	0.83	0.83	1.00
4	32	34	34	33	33	0.94	0.97	1.03
5	23	23	23	22	22	1.00	1.05	1.05
Nivel de Asimetría 3								
3	77	69	68	68	68	1.13	1.13	1.00
4	51	42	42	41	41	1.21	1.24	1.02
5	36	34	34	31	31	1.06	1.16	1.10
Nivel de Asimetría 4								
3	65	64	64	62	62	1.02	1.05	1.03
4	45	41	41	39	39	1.10	1.15	1.05
5	36	28	28	32	26	1.29	1.38	1.08



Cuadro A.15. Tamaños de muestras mínimos para obtener un $CV=0.05$ en la estimación de la media en poblaciones de tamaño 1000 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	15	25	25	24	24	0.60	0.63	1.04
4	11	14	14	12	12	0.79	0.92	1.17
5	9	10	10	7	7	0.90	1.29	1.43
Nivel de Asimetría 2								
3	54	73	73	73	73	0.74	0.74	1.00
4	35	40	40	39	39	0.88	0.90	1.03
5	24	25	25	24	24	0.96	1.00	1.04
Nivel de Asimetría 3								
3	107	96	96	96	96	1.11	1.11	1.00
4	66	60	60	59	59	1.10	1.12	1.02
5	43	41	41	.	41	1.05	1.05	1.00
Nivel de Asimetría 4								
3	98	96	96	96	96	1.02	1.02	1.00
4	60	56	56	54	54	1.07	1.11	1.04
5	44	34	34	31	32	1.29	1.38	1.06



Cuadro A.16. Tamaños de muestras mínimos para obtener un $CV=0.10$ en la estimación de la media en poblaciones de tamaño 200 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	4	8	8	7	7	0.50	0.57	1.14
4	4	5	5	5	5	0.80	0.80	1.00
5	5	6	6	6	6	0.83	0.83	1.00
Nivel de Asimetría 2								
3	15	18	18	17	17	0.83	0.88	1.06
4	10	11	11	9	9	0.91	1.11	1.22
5	9	9	9	7	7	1.00	1.29	1.29
Nivel de Asimetría 3								
3	27	22	22	21	21	1.23	1.29	1.05
4	17	15	15	12	12	1.13	1.42	1.25
5	13	11	11	10	10	1.18	1.30	1.10
Nivel de Asimetría 4								
3	22	21	21	20	20	1.05	1.10	1.05
4	14	13	13	11	11	1.08	1.27	1.18
5	11	9	9	.	7	1.22	1.57	1.29



Cuadro A.17. Tamaños de muestras mínimos para obtener un $CV=0.10$ en la estimación de la media en poblaciones de tamaño 250 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	4	8	8	7	7	0.50	0.57	1.14
4	4	5	5	5	5	0.80	0.80	1.00
5	5	6	6	6	6	0.83	0.83	1.00
Nivel de Asimetría 2								
3	15	19	19	18	18	0.79	0.83	1.06
4	10	11	11	9	9	0.91	1.11	1.22
5	9	9	9	7	7	1.00	1.29	1.29
Nivel de Asimetría 3								
3	29	24	24	24	24	1.21	1.21	1.00
4	18	16	16	14	14	1.13	1.29	1.14
5	13	11	11	11	11	1.18	1.18	1.00
Nivel de Asimetría 4								
3	25	24	24	23	23	1.04	1.09	1.04
4	16	15	15	13	13	1.07	1.23	1.15
5	12	10	10	8	8	1.20	1.50	1.25



Cuadro A.18. Tamaños de muestras mínimos para obtener un $CV=0.10$ en la estimación de la media en poblaciones de tamaño 333 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	4	8	8	7	7	0.50	0.57	1.14
4	4	5	5	5	5	0.80	0.80	1.00
5	5	6	6	6	6	0.83	0.83	1.00
Nivel de Asimetría 2								
3	15	21	21	20	20	0.71	0.75	1.05
4	10	12	12	11	11	0.83	0.91	1.09
5	9	9	9	7	7	1.00	1.29	1.29
Nivel de Asimetría 3								
3	32	28	28	27	27	1.14	1.19	1.04
4	19	18	18	17	17	1.06	1.12	1.06
5	14	13	13	12	12	1.08	1.17	1.08
Nivel de Asimetría 4								
3	30	29	29	27	27	1.03	1.11	1.07
4	18	17	17	15	15	1.06	1.20	1.13
5	13	11	11	8	8	1.18	1.63	1.38



Cuadro A.19. Tamaños de muestras mínimos para obtener un $CV=0.10$ en la estimación de la media en poblaciones de tamaño 500 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	5	8	8	7	7	0.63	0.71	1.14
4	5	5	5	5	5	1.00	1.00	1.00
5	5	6	6	6	6	0.83	0.83	1.00
Nivel de Asimetría 2								
3	15	22	22	21	21	0.68	0.71	1.05
4	10	12	12	12	12	0.83	0.83	1.00
5	9	10	10	7	7	0.90	1.29	1.43
Nivel de Asimetría 3								
3	35	33	33	33	33	1.06	1.06	1.00
4	20	21	21	20	20	0.95	1.00	1.05
5	14	16	16	14	14	0.88	1.00	1.14
Nivel de Asimetría 4								
3	35	34	34	33	33	1.03	1.06	1.03
4	22	19	19	18	18	1.16	1.22	1.06
5	15	12	12	11	11	1.25	1.36	1.09



Cuadro A.20. Tamaños de muestras mínimos para obtener un $CV=0.10$ en la estimación de la media en poblaciones de tamaño 1000 de distintos niveles de asimetría para los cinco métodos, GEO, K, KG, LH, LHG, y eficiencias relativas entre los métodos GEO y KG, GEO y LHG, KG y LHG

Número de estratos	n_{GEO}	n_K	n_{KG}	n_{LH}	n_{LHG}	$efrel_{GEO,KG}$	$efrel_{GEO,LHG}$	$efrel_{KG,LHG}$
Nivel de Asimetría 1								
3	5	9	9	7	7	0.56	0.71	1.29
4	5	5	5	5	5	1.00	1.00	1.00
5	5	6	6	6	6	0.83	0.83	1.00
Nivel de Asimetría 2								
3	15	24	24	23	23	0.63	0.65	1.04
4	11	12	12	12	12	0.92	0.92	1.00
5	9	10	10	8	7	0.90	1.29	1.43
Nivel de Asimetría 3								
3	39	44	44	44	44	0.89	0.89	1.00
4	23	29	29	27	27	0.79	0.85	1.07
5	15	17	17	15	16	0.88	0.94	1.06
Nivel de Asimetría 4								
3	41	42	42	42	42	0.98	0.98	1.00
4	24	23	23	23	23	1.04	1.04	1.00
5	17	15	15	12	13	1.13	1.31	1.15

